STIC Search Report
Biotech-Chem Library

Scientific & Technical Information Center
USPTO

STIC Database Tracking Number: 107433

TO: Janet Epps
Location: cm1/11e01/11e12
Art Unit: 1635
Monday, November 10, 2003
Case Serial Number: 09/817387

From: Paul Schulwitz
Location: Biotech-Chem Library
CM1-6B06
Phone: 305-1954

paul.schulwitz@uspto.gov

Search Notes

Examiner Epps,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(703)305-1954

sira
Search and Information Resources Administration

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into to two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29537
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-29537

Query Match 72.2%; Score 16.6; DB 12; Length 65;
Best Local Similarity 69.6%; Pred. No. 1.6e+02;
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              ||| | |||||||:: ||| :||
Db         47 GTAGTCCTCAGAGTTTGGGGTAG 25
```

RESULT 15
US-09-817-387-2
; Sequence 2, Application US/09817387
; Patent No. US20010039263A1
; GENERAL INFORMATION:
; APPLICANT: Max-Delbruck-Centrum fur Molekulare Medizin
; TITLE OF INVENTION: Chimeric Oligonucleotides and the Use Thereof
; FILE REFERENCE: 101195-24
; CURRENT APPLICATION NUMBER: US/09/817,387
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: DE 197 20 151.2
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide, linkages between positions 1 to
; OTHER INFORMATION: 20 are phosphorothioates, linkages between
; OTHER INFORMATION: positions 20 to 25 are phosphodiesters
US-09-817-387-2

Query Match 67.0%; Score 15.4; DB 9; Length 25;
Best Local Similarity 82.4%; Pred. No. 5.9e+02;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GTACTGCTCAGAGUUAG 17
              | ||||||||||::||
Db          9 GGACTGCTCAGAGTTAG 25
```

Search completed: November 8, 2003, 04:56:13
Job time : 172 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 8, 2003, 04:01:47 ; Search time 170 Seconds
(without alignments)
365.218 Million cell updates/sec

Title: US-09-817-387-16
Perfect score: 23
Sequence: 1 gtactgctcagaguuaggguuag 23

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 2989766

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	16.6	72.2	65	24	ABN56789	Mouse spliced tran
c	2	16.6	72.2	163	11	AAQ06654	Feline T-cell lymp
c	3	16.2	70.4	134	23	ABV05840	Human prostate exp
c	4	15.8	68.7	179	21	AAZ42331	Human 5' EST isola
c	5	15.6	67.8	44	18	AAT69647	Telomerase amplifi
	6	15.2	66.1	51	22	AAL30461	Human SNP oligonuc
c	7	15.2	66.1	145	21	AAC26003	Human secreted pro
c	8	15	65.2	18	18	AAT69649	Telomerase competi
c	9	15	65.2	21	18	AAT66719	Telomerase 3' end
c	10	15	65.2	24	19	AAV49634	Telomerase primer
c	11	15	65.2	29	19	AAV05785	Probe 1 for telome
c	12	15	65.2	29	19	AAV05786	Probe 2 for telome
c	13	15	65.2	29	19	AAV05787	Probe 3 for telome
c	14	15	65.2	29	19	AAV05788	Probe 4 for telome
c	15	15	65.2	29	22	AAH45829	Telomere size dete
	16	15	65.2	34	19	AAV05774	TRAP assay primer.
c	17	15	65.2	39	21	AAZ96507	T cell antigen rec
c	18	15	65.2	48	19	AAV49636	Telomerase primer
	19	15	65.2	51	22	AAH40136	Human SNP flanking
	20	15	65.2	62	18	AAT66718	Telomerase activit
	21	15	65.2	62	24	ABK10342	Zinc finger protei
	22	15	65.2	68	24	ABK10341	Zinc finger protei
	23	15	65.2	70	19	AAV49629	Synthetic telomera
	24	14.8	64.3	172	25	ABX27421	Human GDP-mannose
c	25	14.6	63.5	95	20	AAX36847	Human XLIS gene fr
c	26	14.6	63.5	132	22	ABA69705	Human foetal liver
c	27	14.4	62.6	139	24	ABV88601	Human colon cancer
c	28	14.4	62.6	181	25	ABX85127	Corn ear-derived p
c	29	14.4	62.6	192	25	ABX61117	Arabidopsis thalia
	30	14.2	61.7	28	19	AAV40551	Homo sapiens secre
	31	14.2	61.7	53	25	ABZ81709	Probe for DNase I
	32	14.2	61.7	60	20	AAV63551	PCR standard prime
	33	14.2	61.7	71	18	AAV76032	Staphylococcus aur
	34	14.2	61.7	165	24	ABN23030	Human ORFX polynuc
c	35	14.2	61.7	167	20	AAV86237	EST clone AA35. H
c	36	14	60.9	25	22	AAH40135	SNP specific SNPE
c	37	14	60.9	32	19	AAV05777	Probe for T7 promo
c	38	14	60.9	35	22	AAD06032	Yeast cystathionin
c	39	14	60.9	90	24	ABK36358	HIV DNA encoding T
c	40	14	60.9	100	24	ABK36958	HIV subcassette PC
c	41	14	60.9	200	20	AAH86174	Human single nucle
	42	13.8	60.0	36	10	AAN90556	Tissue plasminogen
c	43	13.8	60.0	53	25	ABZ81710	Probe for DNase I
c	44	13.8	60.0	60	24	ABN43817	Human spliced tran
c	45	13.8	60.0	98	25	ABX54057	Bovine EST associa

ALIGNMENTS

RESULT 1
ABN56789/c
ID ABN56789 standard; DNA; 65 BP.
XX
AC ABN56789;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:29537.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which

PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 29537; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 17 A; 18 C; 16 G; 14 T; 0 other;

Query Match 72.2%; Score 16.6; DB 24; Length 65;
Best Local Similarity 69.6%; Pred. No. 1.6e+02;
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              ||| | ||||||:: ||| :||
Db         47 GTAGTCCTCAGAGTTTGGGGTAG 25
```

RESULT 2
AAQ06654/c
ID AAQ06654 standard; DNA; 163 BP.
XX
AC AAQ06654;
XX
DT 26-FEB-1991 (first entry)
XX
DE Feline T-cell lymphotrophic lentivirus of clone 2BYCXL2.
XX
KW Feline T-cell lymphotropic lentivirus; FIV; 2BYCXL2; antibodies;
KW vaccines; ds.
XX
OS Feline T-cell lymphotropic lentivirus 2428 (Pentaluma).
XX
FH Key Location/Qualifiers
FT CDS 2..163
FT /*tag= a
FT /label=FIV
XX
PN WO9013573-A.
XX
PD 15-NOV-1990.
XX
PF 30-APR-1990; 90WO-US02338.
XX
PR 08-DEC-1989; 89US-0447810.
PR 08-MAY-1989; 89US-0348784.
XX
PA (IDEX-) IDEXX CORP.
XX
PI Anderson PR, Oconnor TP, Tonelli QJ;
XX
DR WPI; 1990-361429/48.
DR P-PSDB; AAR08085.
XX
PT Feline T-cell lympho-tropic lentivirus poly-peptide(s) - used for
PT specific detection of FIV antibodies, prodn. of antibodies and in
PT vaccines
XX
PS Disclosure; Fig 5(b); 37pp; English.
XX
CC FIV nucleic acid is useful for prodn. of large amts. of FIV
CC polypeptides, or fragments, and also for the detection of homologous
CC nucleic acids in vivo. The amino acid sequence derived from this
CC sequence shows homology with the envelope gene of equine infectious
CC anemia virus, a lentivirus, immunologically closely related to FIV.
CC Nucleic acid probes derived from the 2BY DNA hybridises to DNA
CC isolated from FIV infected but not uninfected cells.
CC Strain 2BY has been deposited ATCC 67938.
CC See also AAQ06653-55 and AAR08094-96.
XX
SQ Sequence 163 BP; 28 A; 66 C; 37 G; 30 T; 2 other;

Query Match 72.2%; Score 16.6; DB 11; Length 163;
Best Local Similarity 65.2%; Pred. No. 1.8e+02;
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              ||| ||| || |::||||::||
Db        118 GTAGGGCTTAGGGTTAGGGTTAG 96
```

RESULT 3
ABV05840/c
ID ABV05840 standard; cDNA; 134 BP.
XX
AC ABV05840;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 5831.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 972; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 134 BP; 27 A; 34 C; 25 G; 43 T; 5 other;

Query Match 70.4%; Score 16.2; DB 23; Length 134;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTACTGCTCAGAGUUAGGGUU 21
||||||||| ||| : |||::
Db 96 GTACTGCTCGGAGGTTGGGTT 76

RESULT 4
AAZ42331/c
ID AAZ42331 standard; cDNA; 179 BP.
XX
AC AAZ42331;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST isolated from a cDNA library SEQ ID NO:90.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification; ss.
XX
OS Homo sapiens.
XX
PN WO9953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-038446/03.
DR P-PSDB; AAY64717.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 1; Page 223; 837pp; English.
XX
CC AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 179 BP; 26 A; 67 C; 34 G; 39 T; 13 other;

Query Match 68.7%; Score 15.8; DB 21; Length 179;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ACTGCTCAGAGUUAGGGUUAG 23
||||||||||:: ||: ||
Db 165 ACTGCTCAGAGTKCAGGTGAG 145

RESULT 5
AAT69647/c
ID AAT69647 standard; DNA; 44 BP.
XX
AC AAT69647;
XX
DT 20-FEB-1998 (first entry)
XX
DE Telomerase amplification primer TE-ACT-ST.
XX
KW Telomerase; substrate; primer; detection; 5'-region; retrovirus;
KW long terminal repeat 2; LTR-2; diagnosis; tumour; screening;
KW effector compound; PCR; amplification; ss.
XX
OS Synthetic.
XX
PN DE19644302-A1.
XX
PD 05-JUN-1997.
XX
PF 24-OCT-1996; 96DE-1044302.
XX
PR 28-NOV-1995; 95DE-1044317.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI Emrich T, Hinzpeter M, Karl G, Leying H;
XX
DR WPI; 1997-299542/28.
XX
PT Measuring telomerase activity, useful for tumour diagnosis and
PT compound screening - by extending substrate primer, followed by
PT amplification and immobilising product for detection
XX
PS Example; Page 13; 21pp; German.
XX
CC The present sequence is a telomerase amplification primer, which
CC can be used in a novel method for detecting telomerase activity.
CC The method comprises adding to a test sample a 1st primer, that
CC serves as telomerase substrate, and nucleoside triphosphate (dNTP)
CC and incubating to allow primer extension by the telomerase,
CC amplifying the extension product, immobilising the amplification
CC product (AP) on a solid phase and qualitative and/or quantitative
CC detection of AP, where the substrate primer is preferably from the
CC 5'-region of the long terminal repeat 2 (LTR-2) sequence of a
CC retrovirus. The method can be used to diagnose tumours and screen
CC compounds for effector activity. Immobilisation of AP provides a
CC signal that is reproducibly representative of telomerase activity,
CC eliminates the need for gel electrophoretic separation and
CC provides high sensitivity. Radioactive labels are not required and

```
CC   the method can be automated for routine use. Specific detection is
CC   achieved by proper choice of hybridisation conditions, without
CC   separation of the telomerase extension product. A specific signal
CC   is generated by 1-10 cell equivalents, but for tumour analysis
CC   10-1000 ng of tissue is usually used.
XX
SQ   Sequence 44 BP; 13 A; 16 C; 9 G; 6 T; 0 other;

  Query Match             67.8%;  Score 15.6;  DB 18;  Length 44;
  Best Local Similarity   63.6%;  Pred. No. 4.8e+02;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 TACTGCTCAGAGUUAGGGUUAG 23
              | ||| | || |::||||::||
Db         28 TTCTGGTTAGGGTTAGGGTTAG 7


RESULT 6
AAL30461
ID   AAL30461 standard; DNA; 51 BP.
XX
AC   AAL30461;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #3669.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2440; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 14 A; 8 C; 13 G; 16 T; 0 other;

  Query Match             66.1%;  Score 15.2;  DB 22;  Length 51;
  Best Local Similarity   70.0%;  Pred. No. 7.6e+02;
  Matches   14;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 CTGCTCAGAGUUAGGGUUAG 23
              ||| ||||| ::|||| || 
Db         25 CTGTTCAGATTTAGGGTGAG 44


RESULT 7
AAC26003/c
ID   AAC26003 standard; cDNA; 145 BP.
XX
AC   AAC26003;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 30078.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 30078; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 145 BP; 39 A; 33 C; 19 G; 54 T; 0 other;

  Query Match             66.1%;  Score 15.2;  DB 21;  Length 145;
  Best Local Similarity   75.0%;  Pred. No. 8.5e+02;
  Matches   15;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 ACTGCTCAGAGUUAGGGUUA 22
              |||| ||||||: ||||: |
Db        124 ACTGGTCAGAGTCAGGGTGA 105
```

```
RESULT 8
AAT69649/c
ID   AAT69649 standard; DNA; 18 BP.
XX
AC   AAT69649;
XX
DT   20-FEB-1998  (first entry)
XX
DE   Telomerase competitor oligonucleotide.
XX
KW   Telomerase; substrate; primer; detection; 5'-region; retrovirus;
KW   long terminal repeat 2; LTR-2; diagnosis; tumour; screening;
KW   effector compound; PCR; competitor oligonucleotide; ss.
XX
OS   Synthetic.
XX
PN   DE19644302-A1.
XX
PD   05-JUN-1997.
XX
PF   24-OCT-1996;   96DE-1044302.
XX
PR   28-NOV-1995;   95DE-1044317.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Emrich T,  Hinzpeter M,  Karl G,  Leying H;
XX
DR   WPI; 1997-299542/28.
XX
PT   Measuring telomerase activity, useful for tumour diagnosis and
PT   compound screening  - by extending substrate primer, followed by
PT   amplification and immobilising product for detection
XX
PS   Example; Page 13; 21pp; German.
XX
CC   The present sequence is a telomerase competitor oligonucleotide,
CC   which can be used in a novel method for detecting telomerase
CC   activity. The method comprises adding to a test sample a 1st
CC   primer, that serves as telomerase substrate, and nucleoside
CC   triphosphate (dNTP) and incubating to allow primer extension by the
CC   telomerase, amplifying the extension product, immobilising the
CC   amplification product (AP) on a solid phase and qualitative and/or
CC   quantitative detection of AP, where the substrate primer is
CC   preferably from the 5'-region of the long terminal repeat 2 (LTR-2)
CC   sequence of a retrovirus. The method can be used to diagnose
CC   tumours and screen compounds for effector activity. Immobilisation
CC   of AP provides a signal that is reproducibly representative of
CC   telomerase activity, eliminates the need for gel electrophoretic
CC   separation and provides high sensitivity. Radioactive labels are
CC   not required and the method can be automated for routine use.
CC   Specific detection is achieved by proper choice of hybridisation
CC   conditions, without separation of the telomerase extension product.
CC   A specific signal is generated by 1-10 cell equivalents, but for
CC   tumour analysis 10-1000 ng of tissue is usually used.
XX
SQ   Sequence 18 BP; 4 A; 8 C; 1 G; 5 T; 0 other;

  Query Match             65.2%;  Score 15;  DB 18;  Length 18;
  Best Local Similarity   73.3%;  Pred. No. 8.4e+02;
  Matches   11;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         16 CAGAGTTAGGGTTAG 2


RESULT 9
AAT66719/c
ID   AAT66719 standard; DNA; 21 BP.
XX
AC   AAT66719;
XX
DT   17-DEC-1997  (first entry)
XX
DE   Telomerase 3' end and telomeric repeat junction probe.
XX
KW   Telomerase activity; cancer; human; somatic cell; infertility;
KW   foetal cell; maternal blood; bone marrow; proliferation; protozoa;
KW   fungal infection; ss.
XX
OS   Synthetic.
XX
PN   WO9715687-A1.
XX
PD   01-MAY-1997.
XX
PF   07-JUN-1996;   96WO-US09669.
XX
PR   15-APR-1996;   96US-0632662.
PR   07-JUN-1995;   95US-0482132.
PR   12-APR-1996;   96US-0631554.
XX
PA   (GERO-) GERON CORP.
XX
PI   Harley CB,  Kim NW,  Weinrich SL;
XX
DR   WPI; 1997-259038/23.
XX
PT   Testing cell sample for telomerase activity - by incubating with
PT   substrate to form extended product which is replicated and detected,
PT   useful for cancer prognosis, diagnosis and monitoring
XX
PS   Disclosure; Page 25; 96pp; English.
XX
CC   A method has been developed for testing a cell sample for telomerase
CC   (TM) activity. The method involves: (a) incubating the cell, or its
CC   extract, with a TM substrate so that it is extended by addition of
CC   telomeric repeats; (b) replicating the extended substrate; and (c)
CC   correlating the presence or absence of TM activity with presence or
CC   absence of the extended substrate. The present sequence represents a
CC   probe corresponding to the junction between the 3' end of telomerase
CC   and the telomeric repeats of the telomerase product. The method can be
CC   used to test for elevated levels of TM in human somatic cells, i.e. in
CC   the diagnosis, prognosis and monitoring of cancer. Also low levels of TM
CC   are associated with infertility, TM may indicate foetal cells in
CC   maternal blood and TM can be a marker for bone marrow proliferation and
CC   infection by protozoa or fungi. The method can also be used to identify,
CC   screen and design telomerase inhibitors. The method is simple,
CC   inexpensive, suitable for automation to provide a high throughput system
CC   and all the steps can be carried out in a single reaction vessel.
XX
SQ   Sequence 21 BP; 6 A; 9 C; 1 G; 5 T; 0 other;

  Query Match             65.2%;  Score 15;  DB 18;  Length 21;
  Best Local Similarity   73.3%;  Pred. No. 8.6e+02;
  Matches   11;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         19 CAGAGTTAGGGTTAG 5


RESULT 10
AAV49634/c
ID   AAV49634 standard; DNA; 24 BP.
XX
AC   AAV49634;
XX
DT   21-OCT-1998  (first entry)
XX
DE   Telomerase primer OLS.
XX
KW   Telomerase; detection; hybridisation; sensitivity; primer; ss.
```

```
XX
OS   Synthetic.
XX
PN   DE19705071-A1.
XX
PD   13-AUG-1998.
XX
PF   11-FEB-1997;   97DE-1005071.
XX
PR   11-FEB-1997;   97DE-1005071.
XX
PA   (HEID/) HEIDORN K.
PA   (KRUP/) KRUPP G.
PA   (PARW/) PARWARESCH R.
XX
PI   Heidorn K,  Krupp G,  Parwaresch R;
XX
DR   WPI; 1998-438335/38.
XX
PT   Determination of telomerase by primer extension and amplification -
PT   using reverse primer unable to generate long products following
PT   artificial dimer formation, increases sensitivity and eliminates
PT   false positives
XX
PS   Example; Fig 4; 10pp; German.
XX
CC   AAV49627-V49636 are primers used in a method for detecting telomerase
CC   with high reliability and very low detection limit. The method involves
CC   a reverse primer that has an internal sequence designed so that long
CC   products can not be formed by misplaced hybridisation from the
CC   artificially formed primer dimer. Use of the specified reverse primer
CC   avoids false positives and increases sensitivity, particularly allowing
CC   detection of telomerase in single cells and reliable differentiation
CC   between negative controls and weakly positive samples. The method can be
CC   analysed in real time, i.e. detection is made during the polymerase
CC   chain reaction (PCR), eliminating the need for a separation step by gel
CC   electrophoresis. This results in a very rapid, easily automated and
CC   quantifiable process, and in some embodiments detection can be with a
CC   simple fluorescent photometer.
XX
SQ   Sequence 24 BP; 6 A; 12 C; 1 G; 5 T; 0 other;

  Query Match             65.2%;  Score 15;  DB 19;  Length 24;
  Best Local Similarity   73.3%;  Pred. No. 8.7e+02;
  Matches   11;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         23 CAGAGTTAGGGTTAG 9


RESULT 11
AAV05785/c
ID   AAV05785 standard; DNA; 29 BP.
XX
AC   AAV05785;
XX
DT   19-JUN-1998  (first entry)
XX
DE   Probe 1 for telomerase.
XX
KW   PCR primer; telomerase; telomerase activity detection; cancer cell;
KW   diagnosis; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   14..15
FT                   /*tag=  a
FT                   /note= "backbone modified with acridinium ester"
XX
PN   WO9800563-A1.
XX
PD   08-JAN-1998.
XX
PF   27-JUN-1997;   97WO-JP02251.
XX
PR   28-JUN-1996;   96JP-0169920.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Hashimoto J,  Hirose M,  Yoshimura T;
XX
DR   WPI; 1998-086987/08.
XX
PT   Detection of telomerase activity in cells for cancer diagnosis - by
PT   telomeric repeat amplification protocol followed by hybridisation
PT   protection assay using a non-radioactive (chemoluminescent) label
XX
PS   Disclosure; Page 26; 58pp; Japanese.
XX
CC   This sequence represents a probe used in the method
CC   of the invention. The method is for the detection of telomerase activity
CC   in cells, by: (1) elongating a telomerase substrate (TS primer) using the
CC   cell telomerase; amplifying the elongated substrate by polymerase chain
CC   reaction (PCR) using a second primer (CX primer); (2) hybridising the
CC   amplification product with a probe labelled with a non-radioactive
CC   (preferably chemoluminescent) label; and (3) assaying the label.
CC   Alternatively the TS primer is attached to a promoter sequence (e.g. T7
CC   RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase promoter) and the
CC   elongated substrate is amplified using an RNA polymerase to synthesise
CC   multiple RNA copies of the sequence and a reverse transcriptase to form
CC   DNA copies. Detection of cancer cells and diagnosis of cancer by
CC   detection of telomerase activity in the cells. This may be by invasive
CC   methods (e.g. biopsy of tissue from bladder, uterus, cervix, spleen,
CC   liver, mammary, colon, stomach, lung, kidney, skin,  oesophagus, brain,
CC   mouth etc) or non-invasive methods (e.g. examination of biological
CC   samples such as urine, uterine smear, bladder washings, mouth washings,
CC   colonic washings, duodenal secretion, saliva, sputum, etc). The method is
CC   rapid and of high sensitivity, and does not require the use of a
CC   radioactive label.
XX
SQ   Sequence 29 BP; 6 A; 14 C; 2 G; 7 T; 0 other;

  Query Match             65.2%;  Score 15;  DB 19;  Length 29;
  Best Local Similarity   73.3%;  Pred. No. 8.9e+02;
  Matches   11;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         23 CAGAGTTAGGGTTAG 9


RESULT 12
AAV05786/c
ID   AAV05786 standard; DNA; 29 BP.
XX
AC   AAV05786;
XX
DT   19-JUN-1998  (first entry)
XX
DE   Probe 2 for telomerase.
XX
KW   PCR primer; telomerase; telomerase activity detection; cancer cell;
KW   diagnosis; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   15..16
FT                   /*tag=  a
FT                   /note= "backbone modified with acridinium ester"
XX
PN   WO9800563-A1.
XX
PD   08-JAN-1998.
```

```
XX
PF   27-JUN-1997;   97WO-JP02251.
XX
PR   28-JUN-1996;   96JP-0169920.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Hashimoto J,  Hirose M,  Yoshimura T;
XX
DR   WPI; 1998-086987/08.
XX
PT   Detection of telomerase activity in cells for cancer diagnosis - by
PT   telomeric repeat amplification protocol followed by hybridisation
PT   protection assay using a non-radioactive (chemoluminescent) label
XX
PS   Disclosure; Page 26; 58pp; Japanese.
XX
CC   This sequence represents a probe used in the method
CC   of the invention. The method is for the detection of telomerase activity
CC   in cells, by: (1) elongating a telomerase substrate (TS primer) using the
CC   cell telomerase; amplifying the elongated substrate by polymerase chain
CC   reaction (PCR) using a second primer (CX primer); (2) hybridising the
CC   amplification product with a probe labelled with a non-radioactive
CC   (preferably chemoluminescent) label; and (3) assaying the label.
CC   Alternatively the TS primer is attached to a promoter sequence (e.g. T7
CC   RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase promoter) and the
CC   elongated substrate is amplified using an RNA polymerase to synthesise
CC   multiple RNA copies of the sequence and a reverse transcriptase to form
CC   DNA copies. Detection of cancer cells and diagnosis of cancer by
CC   detection of telomerase activity in the cells. This may be by invasive
CC   methods (e.g. biopsy of tissue from bladder, uterus, cervix, spleen,
CC   liver, mammary, colon, stomach, lung, kidney, skin,  oesophagus, brain,
CC   mouth etc) or non-invasive methods (e.g. examination of biological
CC   samples such as urine, uterine smear, bladder washings, mouth washings,
CC   colonic washings, duodenal secretion, saliva, sputum, etc). The method is
CC   rapid and of high sensitivity, and does not require the use of a
CC   radioactive label.
XX
SQ   Sequence 29 BP; 6 A; 14 C; 2 G; 7 T; 0 other;

  Query Match             65.2%;  Score 15;  DB 19;  Length 29;
  Best Local Similarity   73.3%;  Pred. No. 8.9e+02;
  Matches   11;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         23 CAGAGTTAGGGTTAG 9


RESULT 13
AAV05787/c
ID   AAV05787 standard; DNA; 29 BP.
XX
AC   AAV05787;
XX
DT   19-JUN-1998  (first entry)
XX
DE   Probe 3 for telomerase.
XX
KW   PCR primer; telomerase; telomerase activity detection; cancer cell;
KW   diagnosis; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   9..10
FT                   /*tag=  a
FT                   /note= "backbone modified with acridinium ester"
XX
PN   WO9800563-A1.
XX
PD   08-JAN-1998.
XX
PF   27-JUN-1997;   97WO-JP02251.
XX
PR   28-JUN-1996;   96JP-0169920.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Hashimoto J,  Hirose M,  Yoshimura T;
XX
DR   WPI; 1998-086987/08.
XX
PT   Detection of telomerase activity in cells for cancer diagnosis - by
PT   telomeric repeat amplification protocol followed by hybridisation
PT   protection assay using a non-radioactive (chemoluminescent) label
XX
PS   Disclosure; Page 27; 58pp; Japanese.
XX
CC   This sequence represents a probe used in the method
CC   of the invention. The method is for the detection of telomerase activity
CC   in cells, by: (1) elongating a telomerase substrate (TS primer) using the
CC   cell telomerase; amplifying the elongated substrate by polymerase chain
CC   reaction (PCR) using a second primer (CX primer); (2) hybridising the
CC   amplification product with a probe labelled with a non-radioactive
CC   (preferably chemoluminescent) label; and (3) assaying the label.
CC   Alternatively the TS primer is attached to a promoter sequence (e.g. T7
CC   RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase promoter) and the
CC   elongated substrate is amplified using an RNA polymerase to synthesise
CC   multiple RNA copies of the sequence and a reverse transcriptase to form
CC   DNA copies. Detection of cancer cells and diagnosis of cancer by
CC   detection of telomerase activity in the cells. This may be by invasive
CC   methods (e.g. biopsy of tissue from bladder, uterus, cervix, spleen,
CC   liver, mammary, colon, stomach, lung, kidney, skin,  oesophagus, brain,
CC   mouth etc) or non-invasive methods (e.g. examination of biological
CC   samples such as urine, uterine smear, bladder washings, mouth washings,
CC   colonic washings, duodenal secretion, saliva, sputum, etc). The method is
CC   rapid and of high sensitivity, and does not require the use of a
CC   radioactive label.
XX
SQ   Sequence 29 BP; 7 A; 11 C; 4 G; 7 T; 0 other;

  Query Match             65.2%;  Score 15;  DB 19;  Length 29;
  Best Local Similarity   73.3%;  Pred. No. 8.9e+02;
  Matches   11;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         18 CAGAGTTAGGGTTAG 4


RESULT 14
AAV05788/c
ID   AAV05788 standard; DNA; 29 BP.
XX
AC   AAV05788;
XX
DT   19-JUN-1998  (first entry)
XX
DE   Probe 4 for telomerase.
XX
KW   PCR primer; telomerase; telomerase activity detection; cancer cell;
KW   diagnosis; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   10..11
FT                   /*tag=  a
FT                   /note= "backbone modified with acridinium ester"
XX
PN   WO9800563-A1.
XX
PD   08-JAN-1998.
XX
PF   27-JUN-1997;   97WO-JP02251.
```

XX
PR 28-JUN-1996; 96JP-0169920.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Hashimoto J, Hirose M, Yoshimura T;
XX
DR WPI; 1998-086987/08.
XX
PT Detection of telomerase activity in cells for cancer diagnosis - by
PT telomeric repeat amplification protocol followed by hybridisation
PT protection assay using a non-radioactive (chemoluminescent) label
XX
PS Disclosure; Page 27; 58pp; Japanese.
XX
CC This sequence represents a probe used in the method
CC of the invention. The method is for the detection of telomerase activity
CC in cells, by: (1) elongating a telomerase substrate (TS primer) using the
CC cell telomerase; amplifying the elongated substrate by polymerase chain
CC reaction (PCR) using a second primer (CX primer); (2) hybridising the
CC amplification product with a probe labelled with a non-radioactive
CC (preferably chemoluminescent) label; and (3) assaying the label.
CC Alternatively the TS primer is attached to a promoter sequence (e.g. T7
CC RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase promoter) and the
CC elongated substrate is amplified using an RNA polymerase to synthesise
CC multiple RNA copies of the sequence and a reverse transcriptase to form
CC DNA copies. Detection of cancer cells and diagnosis of cancer by
CC detection of telomerase activity in the cells. This may be by invasive
CC methods (e.g. biopsy of tissue from bladder, uterus, cervix, spleen,
CC liver, mammary, colon, stomach, lung, kidney, skin, oesophagus, brain,
CC mouth etc) or non-invasive methods (e.g. examination of biological
CC samples such as urine, uterine smear, bladder washings, mouth washings,
CC colonic washings, duodenal secretion, saliva, sputum, etc). The method is
CC rapid and of high sensitivity, and does not require the use of a
CC radioactive label.
XX
SQ Sequence 29 BP; 7 A; 11 C; 4 G; 7 T; 0 other;

Query Match 65.2%; Score 15; DB 19; Length 29;
Best Local Similarity 73.3%; Pred. No. 8.9e+02;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         18 CAGAGTTAGGGTTAG 4
```

RESULT 15
AAH45829/c
ID AAH45829 standard; DNA; 29 BP.
XX
AC AAH45829;
XX
DT 11-SEP-2001 (first entry)
XX
DE Telomere size determination method related oligonucleotide #2.
XX
KW Telomere size determination; chromosomal DNA; probe; primer;
KW repetitive sequence; tissue aging; cancer progression; ss.
XX
OS Synthetic.
XX
PN JP2001095586-A.
XX
PD 10-APR-2001.
XX
PF 30-SEP-1999; 99JP-0279948.
XX
PR 30-SEP-1999; 99JP-0279948.
XX
PA (IDET/) IDE T.
XX
DR WPI; 2001-360495/38.
XX
PT Determining telomere size useful for investigating aging in tissue and
PT progression of cancer -
XX
PS Example 1; Page 5; 8pp; Japanese.
XX
CC The present invention describes a method for determining the length of
CC telomeres, involving hybridising a chromosomal DNA extracted from a
CC sample and a labeled DNA probe with a sequence complementary to a
CC repetitive telomeric sequence, and measuring the labeled signal of the
CC hybridised probe to give the size of telomere. This can be used to
CC investigate tissue aging and the progression of cancer and in monitoring
CC the prognosis of patients. The present sequence is an oligonucleotide
CC probe described in the exemplification of the invention.
XX
SQ Sequence 29 BP; 7 A; 14 C; 2 G; 6 T; 0 other;

Query Match 65.2%; Score 15; DB 22; Length 29;
Best Local Similarity 73.3%; Pred. No. 8.9e+02;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          9 CAGAGUUAGGGUUAG 23
              |||||::||||::||
Db         23 CAGAGTTAGGGTTAG 9
```

Search completed: November 8, 2003, 04:53:12
Job time : 172 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 8, 2003, 04:47:47 ; Search time 1609 Seconds
(without alignments)
347.423 Million cell updates/sec

Title: US-09-817-387-16
Perfect score: 23
Sequence: 1 gtactgctcagaguuaggguuag 23

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 2751168

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18.2	79.1	123	28	AQ848512	AQ848512 LMAJFV1_1
c	2	18.2	79.1	177	10	AW860790	AW860790 QV0-CT038
c	3	17.8	77.4	183	10	AW859579	AW859579 MR1-CT035
	4	17.8	77.4	185	10	AW945878	AW945878 QV4-EN004
	5	17.2	74.8	128	12	BM417615	BM417615 952020B09
c	6	16.6	72.2	142	29	TA259E01Q	AL488377 T. brucei
c	7	16.4	71.3	98	9	AW638838	AW638838 bl75c02.w
	8	16.2	70.4	105	9	AA932716	AA932716 oo40d07.s
c	9	16.2	70.4	106	9	AA177143	AA177143 nc02b07.s
c	10	16.2	70.4	106	12	BM751182	BM751182 K-EST0027
c	11	16.2	70.4	109	12	BM836523	BM836523 K-EST0112
	12	16.2	70.4	111	9	AW238289	AW238289 xp20d02.x
c	13	16.2	70.4	116	10	AW868506	AW868506 MR1-SN006
c	14	16.2	70.4	116	10	AW868508	AW868508 MR1-SN006
	15	16.2	70.4	120	10	BE089629	BE089629 QV0-BT070
	16	16.2	70.4	120	10	BE709187	BE709187 QV2-HT057
	17	16.2	70.4	122	10	BE164153	BE164153 QV2-HT046
	18	16.2	70.4	128	9	AW265751	AW265751 xq76f06.x
	19	16.2	70.4	128	10	BE084274	BE084274 PM3-BT065
	20	16.2	70.4	130	10	BE708946	BE708946 QV2-HT057
	21	16.2	70.4	134	10	BE079661	BE079661 RC5-BT062
	22	16.2	70.4	134	10	BE079663	BE079663 RC5-BT062
c	23	16.2	70.4	134	10	BE079796	BE079796 RC6-BT062
c	24	16.2	70.4	134	10	BE093348	BE093348 CM2-BT075
	25	16.2	70.4	154	10	BE816793	BE816793 RC2-BN024
	26	16.2	70.4	160	10	AW946380	AW946380 RC2-ET001
c	27	16.2	70.4	164	12	BM751469	BM751469 K-EST0027
	28	16.2	70.4	165	9	AI702380	AI702380 tz66f09.x
	29	16.2	70.4	165	9	AW805075	AW805075 QV1-UM009
	30	16.2	70.4	167	10	AW996105	AW996105 QV3-BN004
	31	16.2	70.4	167	10	BE828247	BE828247 QV0-ET003
	32	16.2	70.4	168	10	BE161974	BE161974 MR0-HT044
c	33	16.2	70.4	169	9	AA194588	AA194588 zq03e04.r
c	34	16.2	70.4	169	9	AA216249	AA216249 hp0992.se
c	35	16.2	70.4	169	10	AW996027	AW996027 QV3-BN004
c	36	16.2	70.4	169	10	BE816807	BE816807 RC2-BN024
c	37	16.2	70.4	170	9	AA094523	AA094523 cp0618.se
	38	16.2	70.4	172	10	BG231090	BG231090 nah77g04.
c	39	16.2	70.4	173	9	AA187870	AA187870 zp74c10.r
	40	16.2	70.4	173	14	T24894	T24894 EST469 Huma
c	41	16.2	70.4	174	10	BE816808	BE816808 RC2-BN024
c	42	16.2	70.4	176	10	BE093337	BE093337 CM2-BT075
c	43	16.2	70.4	177	9	AA247416	AA247416 csg2159.s
	44	16.2	70.4	178	10	BE816768	BE816768 RC2-BN024
	45	16.2	70.4	179	10	AW897877	AW897877 RC3-NN006

ALIGNMENTS

RESULT 1
AQ848512
LOCUS AQ848512 123 bp DNA linear GSS 25-MAY-2001
DEFINITION LMAJFV1_lm10b07.x1 Leishmania major FV1 random genomic library Leishmania major genomic clone LMAJFV1_lm10b07 3' similar to contains Alu repetitive element;contains 2. .68 LHR+TAS-A.1 leishmania repetitive element ;, genomic survey sequence.
ACCESSION AQ848512
VERSION AQ848512.1 GI:6053160
KEYWORDS GSS.
SOURCE Leishmania major
ORGANISM Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
REFERENCE 1 (bases 1 to 123)
AUTHORS Akopyants,N.S., Clifton,S.W., Martin,J., Pape,D., Wylie,T., Li,L., Kissinger,J.C, Roos,D.S. and Beverley,S.M.
TITLE A survey of the Leishmania major Friedlin strain V1 genome by shotgun sequencing: a resource for DNA microarrays and expression profiling
JOURNAL Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)
MEDLINE 21192569
PUBMED 11295190
COMMENT Other_GSSs: lm10b07.y1
Contact: Akopyants, NS / Beverley, SM
WashU Leishmania Project
Washington University School of Medicine

```
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library construction: Natalia S. Akopyants, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             If using this information please cite:
             N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major
            Friedlin strain V1 genome by shotgun sequencing' and the Washington
            University Genome Sequencing Center For information on obtaining
            clone material please contact: Natalia S. Akopyants Ph.D.
            (natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.
            (beverley@borcim.wustl.edu)
            Seq primer: -40UP from Gibco
            Class: shotgun
            High quality sequence stop: 108.
FEATURES              Location/Qualifiers
     source           1. .123
                      /organism="Leishmania major"
                      /mol_type="genomic DNA"
                      /strain="Friedlin strain V1"
                      /db_xref="taxon:5664"
                      /clone="LMAJFV1_lm10b07"
                      /lab_host="TOP10 (Invitrogen)"
                      /clone_lib="Leishmania major FV1 random genomic library"
                      /note="Vector: pZero-2 (Invitrogen); Site_1: EcoRV;
                      Genomic DNA was isolated from stationary phase cells. For
                      this library, DNA was sheared to give a tight size
                      distribution of 1-1.5kb fragments, blunt-ended with T4 DNA
                      polymerase, dephosphorylated with Shrimp Alkaline
                      Phosphatase and ligated into pZero-2 vector's EcoRV site."
BASE COUNT       19 a     24 c     44 g     36 t
ORIGIN

  Query Match             79.1%;  Score 18.2;  DB 28;  Length 123;
  Best Local Similarity   69.6%;  Pred. No. 4.2e+02;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              |||||| | || |::||||::||
Db         79 GTACTGGTTAGGGTTAGGGTTAG 101


RESULT 2
AW860790/c
LOCUS       AW860790                177 bp    mRNA    linear   EST 19-MAY-2000
DEFINITION  QV0-CT0383-210400-206-a12 CT0383 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW860790
VERSION     AW860790.1  GI:7956392
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 177)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV0-CT0383-210
            400-206-a12&t3=2000-04-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 177.
FEATURES              Location/Qualifiers
     source           1. .177
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="CT0383"
                      /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       67 a     42 c     32 g     36 t
ORIGIN

  Query Match             79.1%;  Score 18.2;  DB 10;  Length 177;
  Best Local Similarity   73.9%;  Pred. No. 4.8e+02;
  Matches   17;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              ||||||||| ||| : |||::||
Db         95 GTACTGCTCGGAGGTTGGGTTAG 73


RESULT 3
AW859579/c
LOCUS       AW859579                183 bp    mRNA    linear   EST 19-MAY-2000
DEFINITION  MR1-CT0355-180200-008-a05 CT0355 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW859579
VERSION     AW859579.1  GI:7955272
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 183)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR1-CT0355-180
            200-008-a05&t3=2000-02-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 20
            High quality sequence stop: 183.
FEATURES              Location/Qualifiers
     source           1. .183
```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0355"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 62 a 48 c 36 g 37 t
ORIGIN

Query Match 77.4%; Score 17.8; DB 10; Length 183;
Best Local Similarity 71.4%; Pred. No. 7.5e+02;
Matches 15; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTACTGCTCAGAGUUAGGGUU 21
||||||||| ||||:: |||::
Db 157 GTACTGCTCGGAGTTTGGGTT 137

RESULT 4
AW945878
LOCUS AW945878 185 bp mRNA linear EST 31-MAY-2000
DEFINITION QV4-EN0040-130500-213-c06 EN0040 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW945878
VERSION AW945878.1 GI:8123636
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 185)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV4-EN0040-130500-213-c06&t3=2000-05-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 185.
FEATURES Location/Qualifiers
source 1. .185
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0040"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 43 a 29 c 51 g 62 t
ORIGIN

Query Match 77.4%; Score 17.8; DB 10; Length 185;
Best Local Similarity 76.2%; Pred. No. 7.5e+02;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTACTGCTCAGAGUUAGGGUU 21
||||||||| ||| :|||::
Db 58 GTACTGCTCGGAGGTAGGGTT 78

RESULT 5
BM417615
LOCUS BM417615 128 bp mRNA linear EST 28-JAN-2002
DEFINITION 952020B09.y1 952 - BMS tissue from Walbot Lab (reduced rRNA) Zea mays cDNA, mRNA sequence.
ACCESSION BM417615
VERSION BM417615.1 GI:18384416
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 128)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 952020 row: B column: 09.
FEATURES Location/Qualifiers
source 1. .128
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="BMS (Black Mexican Sweet)"
/db_xref="taxon:4577"
/tissue_type="suspension culture"
/dev_stage="mixed logarithmic and stationary growth phases"
/lab_host="DH10B"
/clone_lib="952 - BMS tissue from Walbot Lab (reduced rRNA)"
/note="Vector: pUC19; Site_1: EcoRI; Site_2: EcoRI; The library was prepared by George Rudenko using poly (A) selected RNA and Universal Riboclone cDNA Synthesis System (Promega). cDNA was synthesized using both random and oligo(dT) primers in separate reactions and equipped with EcoRI adaptors. Library was size-fractionated on agarose gels (for insert size >400bp) and non-directionally cloned into EcoRI-digested pUC19 vector. Blue/white selection on carbenicillin-containing plates was used to recover positive clones."
BASE COUNT 36 a 23 c 28 g 41 t
ORIGIN

Query Match 74.8%; Score 17.2; DB 12; Length 128;
Best Local Similarity 68.2%; Pred. No. 1.2e+03;
Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTACTGCTCAGAGUUAGGGUUA 22
||| |||||||::||||::|
Db 9 GTAGAACTCAGAGTTAGGGTTA 30

```
RESULT 6
TA259E01Q/c
LOCUS       TA259E01Q                142 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 259e01, reverse sequence,
            genomic survey sequence.
ACCESSION   AL488377
VERSION     AL488377.1  GI:11863771
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 142)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="259e01"
BASE COUNT       41 a     69 c      9 g     23 t
ORIGIN

  Query Match             72.2%;  Score 16.6;  DB 29;  Length 142;
  Best Local Similarity   65.2%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              |||| | | || |::||||::||
Db         80 GTACGGGTTAGGGTTAGGGTTAG 58


RESULT 7
AW638838/c
LOCUS       AW638838                 98 bp    mRNA    linear   EST 26-APR-2001
DEFINITION  bl75c02.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
            laevis cDNA clone PBX0075C02 5', mRNA sequence.
ACCESSION   AW638838
VERSION     AW638838.1  GI:7396005
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Blackshear,P.J., Lai,W.S., Thorn,J.M., Kennington,E.A., Staffa,N.G.
            Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M., Touchman
            ,J.W., Bonaldo,M.F. and Soares,M.B.
  TITLE     The NIEHS Xenopus maternal EST project: interim analysis of the
            first 13,879 ESTs from unfertilized eggs
  JOURNAL   Gene 267 (1), 71-87 (2001)
  MEDLINE   21211403
   PUBMED   11311557
COMMENT     Contact: Perry J. Blackshear
            Office of Clinical Research and Laboratory of Signal Transduction
            National Institute of Environmental Health Sciences
            A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
            USA
            Tel: 919 541-4899
            Fax: 919 541-4571
            Email: black009@niehs.nih.gov
            Clone is available through Research Genetics, Inc., 2130 Memorial
            Parkway, Huntsville, AL 35901
             phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
            cdna@resgen.com
             DNA Sequencing and analyses performed by National Institutes of
            Health Intramural Sequencing Center (NISC).
            PCR PRimers
            FORWARD: TGTAAAACGACGGCCAGT
            BACKWARD: CAGGAAACAGCTATGACC
            Plate: 0075  row: C  column: 02
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="PBX0075C02"
                     /sex="female"
                     /tissue_type="unfertilized egg"
                     /cell_type="unfertilized egg"
                     /dev_stage="unfertilized egg"
                     /lab_host="DH10B"
                     /clone_lib="Blackshear/Soares normalized Xenopus egg
                     library"
                     /note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
                     PolyA-selected mRNA was prepared from unfertilized Xenopus
                     laevis eggs.  The library was constructed in the vector
                     pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
                     Soares, M.B. 'Normalization and subtraction:  two
                     approaches to facilitate gene discovery',  Genome Research
                     6:791-806, 1996.  The first strand synthesis used a
                     NotI-dT18 primer; double stranded cDNAs were ligated to
                     EcoRI adapters, digested with NotI, and directionally
                     cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
                     The library contained approximately 7.2 X 10^5
                     recombinants, with average insert sizes of 1-1.5 kb."
BASE COUNT       27 a     27 c     25 g     19 t
ORIGIN

  Query Match             71.3%;  Score 16.4;  DB 9;  Length 98;
  Best Local Similarity   72.2%;  Pred. No. 2.5e+03;
  Matches   13;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 GCTCAGAGUUAGGGUUAG 23
              |||||| |::||||::||
Db         96 GCTCAGGGTTAGGGTTAG 79


RESULT 8
AA932716
LOCUS       AA932716                 105 bp    mRNA    linear   EST 24-APR-1998
DEFINITION  oo40d07.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1568653 3',
            mRNA sequence.
ACCESSION   AA932716
VERSION     AA932716.1  GI:3086681
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 105)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 80.
FEATURES              Location/Qualifiers
     source           1. .105
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1568653"
                      /tissue_type="carcinoid"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Lu5"
                      /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; 1st strand cDNA was prepared from
                      neuroendocrine lung carcinoid, and was then primed with a
                      Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                      to Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector.  Library is normalized. Library was
                      constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       19 a     20 c     30 g     36 t
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 9;  Length 105;
  Best Local Similarity   71.4%;  Pred. No. 3.1e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         36 GTACTGCTCGGAGGTTGGGTT 56


RESULT 9
AA177143/c
LOCUS       AA177143                106 bp    mRNA    linear   EST 26-AUG-1998
DEFINITION  nc02b07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:211, mRNA
            sequence.
ACCESSION   AA177143
VERSION     AA177143.1  GI:1758301
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 76.
FEATURES              Location/Qualifiers
     source           1. .106
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="ATCC (inhost):1363284"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:211"
                      /sex="Male"
                      /dev_stage="45 years old"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Pr3"
                      /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                      strand cDNA was primed with oligo(dT)17 on 50 ng of
                      DNAse-treated, total cellular RNA obtained from 5,000-10
                      ,000 microdissected cells histologically-determined to be
                      fully malignant prostate cancer cells.  Double-stranded
                      cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                      applied to the cDNA with an adaptor-specific primer, and
                      the resulting PCR product subcloned into pAMP10 by the
                      UDG-cloning method (Life Technologies).  Average insert
                      size is 600 bp. NOTE: Not directionally cloned.  This
                      library was constructed by  David Krizman."
BASE COUNT       34 a     33 c     21 g     17 t      1 others
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 9;  Length 106;
  Best Local Similarity   71.4%;  Pred. No. 3.1e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         96 GTACTGCTCGGAGGTTGGGTT 76


RESULT 10
BM751182/c
LOCUS       BM751182                106 bp    mRNA    linear   EST 04-MAR-2002
DEFINITION  K-EST0027207 S9SNU601 Homo sapiens cDNA clone S9SNU601-11-D05 5',
            mRNA sequence.
ACCESSION   BM751182
VERSION     BM751182.1  GI:19080800
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 11  row: D  column: 05
            High quality sequence stop: 106.
FEATURES              Location/Qualifiers
     source           1. .106
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="S9SNU601-11-D05"
                      /sex="M"
                      /tissue_type="Ascites"
                      /cell_type="Epithelial"
                      /cell_line="SNU-601"
                      /lab_host="Top10F'"
                      /clone_lib="S9SNU601"
```

/note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI; Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 36 a 31 c 20 g 19 t
ORIGIN

Query Match 70.4%; Score 16.2; DB 12; Length 106;
Best Local Similarity 71.4%; Pred. No. 3.1e+03;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTACTGCTCAGAGUUAGGGUU 21
||||||||| ||| : |||::
Db 70 GTACTGCTCGGAGGTTGGGTT 50

RESULT 11
BM836523/c
LOCUS BM836523 109 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0112212 S9SNU601 Homo sapiens cDNA clone S9SNU601-63-E06 5', mRNA sequence.
ACCESSION BM836523
VERSION BM836523.1 GI:19192932
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: E column: 06
High quality sequence stop: 109.
FEATURES Location/Qualifiers
source 1. .109
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-63-E06"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F'"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI; Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 38 a 35 c 17 g 19 t
ORIGIN

Query Match 70.4%; Score 16.2; DB 12; Length 109;
Best Local Similarity 71.4%; Pred. No. 3.2e+03;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTACTGCTCAGAGUUAGGGUU 21
||||||||| ||| : |||::
Db 32 GTACTGCTCGGAGGTTGGGTT 12

RESULT 12
AW238289
LOCUS AW238289 111 bp mRNA linear EST 13-DEC-1999
DEFINITION xp20d02.x1 NCI_CGAP_HN10 Homo sapiens cDNA clone IMAGE:2740899 3', mRNA sequence.
ACCESSION AW238289
VERSION AW238289.1 GI:6570606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 111)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D., Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 104.
FEATURES Location/Qualifiers
source 1. .111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2740899"
/tissue_type="carcinoma in situ from retromolar trigone"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_HN10"
/note="Vector: pAMP10; cDNA made by oligo-dT priming. Non-directionally cloned into the UDG sites of pAMP10. Size-selected on agarose gel, average insert size 500 bp. Primary library; non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D (NCI). Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."
BASE COUNT 19 a 20 c 33 g 39 t
ORIGIN

Query Match 70.4%; Score 16.2; DB 9; Length 111;
Best Local Similarity 71.4%; Pred. No. 3.2e+03;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         80 GTACTGCTCGGAGGTTGGGTT 100


RESULT 13
AW868506/c
LOCUS       AW868506                116 bp    mRNA    linear   EST 22-MAY-2000
DEFINITION  MR1-SN0063-040500-001-a03_1 SN0063 Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   AW868506
VERSION     AW868506.1  GI:8002545
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 116)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR1-SN0063-040
            500-001-a03_1&t3=2000-05-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 116.
FEATURES             Location/Qualifiers
     source          1. .116
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="SN0063"
                     /note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       37 a     36 c     22 g     21 t
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 10;  Length 116;
  Best Local Similarity   71.4%;  Pred. No. 3.3e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         83 GTACTGCTCGGAGGTTGGGTT 63


RESULT 14
AW868508/c
LOCUS       AW868508                116 bp    mRNA    linear   EST 22-MAY-2000
DEFINITION  MR1-SN0063-040500-001-b03_1 SN0063 Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   AW868508
VERSION     AW868508.1  GI:8002547
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 116)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR1-SN0063-040
            500-001-b03_1&t3=2000-05-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 116.
FEATURES             Location/Qualifiers
     source          1. .116
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="SN0063"
                     /note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       37 a     36 c     22 g     21 t
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 10;  Length 116;
  Best Local Similarity   71.4%;  Pred. No. 3.3e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         83 GTACTGCTCGGAGGTTGGGTT 63


RESULT 15
BE089629
LOCUS       BE089629                120 bp    mRNA    linear   EST 12-JUN-2000
DEFINITION  QV0-BT0703-120500-225-e10 BT0703 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE089629
VERSION     BE089629.1  GI:8480047
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 120)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
```

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV0-BT0703-120500-225-e10&t3=2000-05-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 120.

FEATURES Location/Qualifiers
source 1. .120
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0703"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 23 a 27 c 32 g 38 t

ORIGIN

Query Match 70.4%; Score 16.2; DB 10; Length 120;
Best Local Similarity 71.4%; Pred. No. 3.3e+03;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         67 GTACTGCTCGGAGGTTGGGTT 87
```

Search completed: November 8, 2003, 05:49:03
Job time : 1616 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 8, 2003, 04:09:27 ; Search time 1533 Seconds
(without alignments)
613.778 Million cell updates/sec

Title: US-09-817-387-16
Perfect score: 23
Sequence: 1 gtactgctcagaguuaggguuag 23

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 1812986

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21.4	93.0	35	6	BD084922	BD084922 Chimeric
2		19.8	86.1	31	6	BD084916	BD084916 Chimeric
3		18.8	81.7	31	6	BD084925	BD084925 Chimeric
4		18.2	79.1	31	6	BD084919	BD084919 Chimeric
5	c	18.2	79.1	160	3	AF031203	AF031203 Leishmani
6	c	16.2	70.4	174	9	HS62F7F	Z55760 H.sapiens C
7	c	16.2	70.4	182	9	HS57C11F	Z61706 H.sapiens C
8	c	16.2	70.4	190	9	HS169H8F	Z57252 H.sapiens C
9		15.8	68.7	118	5	GGA240706	AJ240706 Gallus ga
10	c	15.6	67.8	44	6	A79663	A79663 Sequence 12
11	c	15.6	67.8	44	6	AR147337	AR147337 Sequence
12	c	15.6	67.8	196	5	AF427996	AF427996 Fringilla
13		15.4	67.0	25	6	BD084917	BD084917 Chimeric
14		15.4	67.0	25	6	BD084920	BD084920 Chimeric
15	c	15.2	66.1	145	6	BD049748	BD049748 Sequence
16	c	15.2	66.1	176	5	AF427997	AF427997 Fringilla
17	c	15	65.2	18	6	A79665	A79665 Sequence 14
18	c	15	65.2	18	6	AR147339	AR147339 Sequence
19	c	15	65.2	21	6	AR037852	AR037852 Sequence
20	c	15	65.2	21	6	AR069385	AR069385 Sequence
21	c	15	65.2	21	6	AR087788	AR087788 Sequence
22	c	15	65.2	21	6	AR211028	AR211028 Sequence
23	c	15	65.2	29	6	AR257904	AR257904 Sequence
24	c	15	65.2	29	6	AR257905	AR257905 Sequence
25	c	15	65.2	29	6	AR257906	AR257906 Sequence
26	c	15	65.2	29	6	AR257907	AR257907 Sequence
27	c	15	65.2	29	6	E15450	E15450 Oligonucleo
28	c	15	65.2	29	6	E40924	E40924 Method for
29		15	65.2	34	6	AR257913	AR257913 Sequence
30		15	65.2	38	6	AR307306	AR307306 Sequence
31		15	65.2	51	6	AX117809	AX117809 Sequence
32		15	65.2	62	6	AR037864	AR037864 Sequence
33		15	65.2	62	6	AR054745	AR054745 Sequence
34		15	65.2	62	6	AR069397	AR069397 Sequence
35		15	65.2	62	6	AR243525	AR243525 Sequence
36		15	65.2	62	6	AX395632	AX395632 Sequence
37		15	65.2	62	6	BD011314	BD011314 Human tel
38		15	65.2	62	6	E37063	E37063 Human telom
39		15	65.2	68	6	AX395585	AX395585 Sequence
40		15	65.2	76	6	A60813	A60813 Sequence 12
41		15	65.2	125	10	MMU403205	AJ403205 M.musculu
42		15	65.2	128	5	GGA240711	AJ240711 Gallus ga
43		15	65.2	135	5	GGA240728	AJ240728 Gallus ga
44		15	65.2	145	5	GGA240732	AJ240732 Gallus ga
45		15	65.2	148	5	GGA240702	AJ240702 Gallus ga

ALIGNMENTS

RESULT 1
BD084922
LOCUS BD084922 35 bp DNA linear PAT 27-AUG-2002
DEFINITION Chimeric oligonucleotides and the use thereof.
ACCESSION BD084922
VERSION BD084922.1 GI:22630532
KEYWORDS JP 2001524972-A/7.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 35)
AUTHORS Matthes,E. and Lipinski,M.V.J.
TITLE Chimeric oligonucleotides and the use thereof
JOURNAL Patent: JP 2001524972-A 7 04-DEC-2001;
MAX DELBRUCK CENTRUM FUR MOLEKULARE MEDIZIN

```
COMMENT     OS   Artificial Sequence
            PN   JP 2001524972-A/7
            PD   04-DEC-2001
            PF   04-MAY-1998 JP 1998547618
            PR   02-MAY-1997 DE 197 20 151.2
            PI   ECKART MATTHES,MARTIN VON JANTA LIPINSKI
            PC   C07H21/00
            CC   Description of Artificial Sequence:chimera oligonucleotide FH
             Key              Location/Qualifiers
            FT   source          1. .35
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .35
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
BASE COUNT       11 a      5 c      10 g      9 t
ORIGIN

  Query Match           93.0%;  Score 21.4;  DB 6;  Length 35;
  Best Local Similarity   78.3%;  Pred. No. 2.8;
  Matches   18;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              | ||||||||||::||||::||
Db          9 GGACTGCTCAGAGTTAGGGTTAG 31


RESULT 2
BD084916
LOCUS       BD084916                 31 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Chimeric oligonucleotides and the use thereof.
ACCESSION   BD084916
VERSION     BD084916.1  GI:22630526
KEYWORDS    JP 2001524972-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Matthes,E. and Lipinski,M.V.J.
  TITLE     Chimeric oligonucleotides and the use thereof
  JOURNAL   Patent: JP 2001524972-A 1 04-DEC-2001;
            MAX DELBRUCK CENTRUM FUR MOLEKULARE MEDIZIN
COMMENT     OS   Artificial Sequence
            PN   JP 2001524972-A/1
            PD   04-DEC-2001
            PF   04-MAY-1998 JP 1998547618
            PR   02-MAY-1997 DE 197 20 151.2
            PI   ECKART MATTHES,MARTIN VON JANTA LIPINSKI
            PC   C07H21/00
            CC   Description of Artificial Sequence:chimera oligonucleotide FH
             Key              Location/Qualifiers
            FT   source          1. .31
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .31
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
BASE COUNT        8 a      4 c      9 g     10 t
ORIGIN

  Query Match           86.1%;  Score 19.8;  DB 6;  Length 31;
  Best Local Similarity   73.9%;  Pred. No. 21;
  Matches   17;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              |||||  ||||||::||||::||
Db          9 GTACTCGTCAGAGTTAGGGTTAG 31


RESULT 3
BD084925
LOCUS       BD084925                 31 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Chimeric oligonucleotides and the use thereof.
ACCESSION   BD084925
VERSION     BD084925.1  GI:22630535
KEYWORDS    JP 2001524972-A/10.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Matthes,E. and Lipinski,M.V.J.
  TITLE     Chimeric oligonucleotides and the use thereof
  JOURNAL   Patent: JP 2001524972-A 10 04-DEC-2001;
            MAX DELBRUCK CENTRUM FUR MOLEKULARE MEDIZIN
COMMENT     OS   Artificial Sequence
            PN   JP 2001524972-A/10
            PD   04-DEC-2001
            PF   04-MAY-1998 JP 1998547618
            PR   02-MAY-1997 DE 197 20 151.2
            PI   ECKART MATTHES,MARTIN VON JANTA LIPINSKI
            PC   C07H21/00
            CC   Description of Artificial Sequence:chimera oligonucleotide CC
             n represents 3' azidodeoxyguanosin
            FH   Key              Location/Qualifiers
            FT   source          1. .31
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .31
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
BASE COUNT        8 a      4 c      8 g     10 t      1 others
ORIGIN

  Query Match           81.7%;  Score 18.8;  DB 6;  Length 31;
  Best Local Similarity   72.7%;  Pred. No. 74;
  Matches   16;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUA 22
              |||||  ||||||::||||::|
Db          9 GTACTCGTCAGAGTTAGGGTTA 30


RESULT 4
BD084919
LOCUS       BD084919                 31 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Chimeric oligonucleotides and the use thereof.
ACCESSION   BD084919
VERSION     BD084919.1  GI:22630529
KEYWORDS    JP 2001524972-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Matthes,E. and Lipinski,M.V.J.
  TITLE     Chimeric oligonucleotides and the use thereof
  JOURNAL   Patent: JP 2001524972-A 4 04-DEC-2001;
            MAX DELBRUCK CENTRUM FUR MOLEKULARE MEDIZIN
COMMENT     OS   Artificial Sequence
            PN   JP 2001524972-A/4
            PD   04-DEC-2001
            PF   04-MAY-1998 JP 1998547618
            PR   02-MAY-1997 DE 197 20 151.2
            PI   ECKART MATTHES,MARTIN VON JANTA LIPINSKI
            PC   C07H21/00
            CC   Description of Artificial Sequence:chimera oligonucleotide FH
             Key              Location/Qualifiers
            FT   source          1. .31
            FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
     source         1. .31
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
```

```
BASE COUNT       9 a      4 c      8 g     10 t
ORIGIN

  Query Match             79.1%;  Score 18.2;  DB 6;  Length 31;
  Best Local Similarity   69.6%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              |||||  ||||||::|| |::||
Db          9 GTACTCGTCAGAGTTAGAGTTAG 31


RESULT 5
AF031203/c
LOCUS       AF031203                160 bp    DNA     linear   INV 20-MAY-1998
DEFINITION  Leishmania major strain 1503 telomere-associated sequence, clone
            major7.
ACCESSION   AF031203
VERSION     AF031203.1  GI:3142346
KEYWORDS    .
SOURCE      Leishmania major
  ORGANISM  Leishmania major
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.
REFERENCE   1  (bases 1 to 160)
  AUTHORS   Fu,G. and Barker,D.C.
  TITLE     Characterisation of Leishmania telomeres reveals unusual telomeric
            repeats and conserved telomere-associated sequence
  JOURNAL   Nucleic Acids Res. 26 (9), 2161-2167 (1998)
  MEDLINE   98213745
   PUBMED   9547275
REFERENCE   2  (bases 1 to 160)
  AUTHORS   Fu,G. and Barker,D.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1997) Department of Pathology, University of
            Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK
FEATURES             Location/Qualifiers
     source          1. .160
                     /organism="Leishmania major"
                     /mol_type="genomic DNA"
                     /strain="1503"
                     /db_xref="taxon:5664"
                     /clone="major7"
     misc_feature    1. .129
                     /note="telomere-associated sequence"
BASE COUNT       49 a     54 c     30 g     27 t
ORIGIN

  Query Match             79.1%;  Score 18.2;  DB 3;  Length 160;
  Best Local Similarity   69.6%;  Pred. No. 1.3e+02;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUUAG 23
              ||||||  | ||  |::||||::||
Db         31 GTACTGGTTAGGGTTAGGGTTAG 9


RESULT 6
HS62F7F/c
LOCUS       HS62F7F                 174 bp    DNA     linear   PRI 17-OCT-1995
DEFINITION  H.sapiens CpG island DNA genomic Mse1 fragment, clone 62f7, forward
            read cpg62f7.ft1a.
ACCESSION   Z55760
VERSION     Z55760.1  GI:1021801
KEYWORDS    CpG island; genomic Mse1 fragment.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE     Purification of CpG islands using a methylated DNA binding column
  JOURNAL   Nat. Genet. 6 (3), 236-244 (1994)
  MEDLINE   94282070
   PUBMED   8012384
REFERENCE   2  (bases 1 to 174)
  AUTHORS   MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT     Vector: pGEM-5Zf(-)
            Clones are available from the UK MRC Human Genome Mapping Project
            Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
            http://www.hgmp.mrc.ac.uk/ for details
            or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .174
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="62f7"
                     /sex="male"
                     /tissue_type="blood"
                     /clone_lib="CGI-1"
                     /dev_stage="adult"
BASE COUNT       60 a     46 c     32 g     36 t
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 9;  Length 174;
  Best Local Similarity   71.4%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              |||||||| ||| : |||::
Db         53 GTACTGCTCGGAGGTTGGGTT 33


RESULT 7
HS57C11F/c
LOCUS       HS57C11F                182 bp    DNA     linear   PRI 22-OCT-1995
DEFINITION  H.sapiens CpG island DNA genomic Mse1 fragment, clone 57c11,
            forward read cpg57c11.ft1a.
ACCESSION   Z61706
VERSION     Z61706.1  GI:1034084
KEYWORDS    CpG island; genomic Mse1 fragment.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE     Purification of CpG islands using a methylated DNA binding column
  JOURNAL   Nat. Genet. 6 (3), 236-244 (1994)
  MEDLINE   94282070
   PUBMED   8012384
REFERENCE   2  (bases 1 to 182)
  AUTHORS   MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT     Vector: pGEM-5Zf(-)
            Clones are available from the UK MRC Human Genome Mapping Project
            Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
            http://www.hgmp.mrc.ac.uk/ for details
            or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .182
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="57c11"
                     /sex="male"
                     /tissue_type="blood"
                     /clone_lib="CGI-1"
                     /dev_stage="adult"
BASE COUNT       61 a     46 c     34 g     39 t      2 others
```

```
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 9;  Length 182;
  Best Local Similarity   71.4%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         53 GTACTGCTCGGAGGTTGGGTT 33


RESULT 8
HS169H8F/c
LOCUS       HS169H8F                190 bp    DNA     linear   PRI 18-OCT-1995
DEFINITION  H.sapiens CpG island DNA genomic Mse1 fragment, clone 169h8,
            forward read cpg169h8.ft1a.
ACCESSION   Z57252
VERSION     Z57252.1  GI:1028483
KEYWORDS    CpG island; genomic Mse1 fragment.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE     Purification of CpG islands using a methylated DNA binding column
  JOURNAL   Nat. Genet. 6 (3), 236-244 (1994)
  MEDLINE   94282070
   PUBMED   8012384
REFERENCE   2  (bases 1 to 190)
  AUTHORS   Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT     Vector: pGEM-5Zf(-)
            Clones are available from the UK MRC Human Genome Mapping Project
            Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
            http://www.hgmp.mrc.ac.uk/ for details
            or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .190
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="169h8"
                     /sex="male"
                     /tissue_type="blood"
                     /clone_lib="CGI-1"
                     /dev_stage="adult"
BASE COUNT       61 a     50 c     37 g     40 t      2 others
ORIGIN

  Query Match             70.4%;  Score 16.2;  DB 9;  Length 190;
  Best Local Similarity   71.4%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAGGGUU 21
              ||||||||| ||| : |||::
Db         53 GTACTGCTCGGAGGTTGGGTT 33


RESULT 9
GGA240706
LOCUS       GGA240706               118 bp    DNA     linear   VRT 05-APR-1999
DEFINITION  Gallus gallus HD7/E2 telomere junction.
ACCESSION   AJ240706
VERSION     AJ240706.1  GI:4583601
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1
  AUTHORS   Thomson,P.A. and Burke,T.
  TITLE     The isolation of chicken telomere junction fragments
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 118)
  AUTHORS   Thomson,P.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-1992) Thomson P.A., Department of Biology,
            University of Leicester, University Road, Leicester, LEI 7RH,
            UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .118
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9031"
                     /clone="HD7/E2"
                     /clone_lib="anchored PCR library"
BASE COUNT       19 a      7 c     49 g     43 t
ORIGIN

  Query Match             68.7%;  Score 15.8;  DB 5;  Length 118;
  Best Local Similarity   68.4%;  Pred. No. 2.8e+03;
  Matches   13;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 TGCTCAGAGUUAGGGUUAG 23
              |||| || |::||||::||
Db         99 TGCTTAGGGTTAGGGTTAG 117


RESULT 10
A79663/c
LOCUS       A79663                  44 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 12 from Patent WO9720069.
ACCESSION   A79663
VERSION     A79663.1  GI:6092617
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Emrich,T. and Leying,H.
  TITLE     METHOD OF DETECTING TELOMERASE ACTIVITY
  JOURNAL   Patent: WO 9720069-A 12 05-JUN-1997;
            BOEHRINGER MANNHEIM GMBH (DE); EMRICH THOMAS (DE)
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
BASE COUNT       13 a     16 c      9 g      6 t
ORIGIN

  Query Match             67.8%;  Score 15.6;  DB 6;  Length 44;
  Best Local Similarity   63.6%;  Pred. No. 4e+03;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 TACTGCTCAGAGUUAGGGUUAG 23
              | ||| | || |::||||::||
Db         28 TTCTGGTTAGGGTTAGGGTTAG 7


RESULT 11
AR147337/c
LOCUS       AR147337                44 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 12 from patent US 6221584.
ACCESSION   AR147337
VERSION     AR147337.1  GI:15111140
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Emrich,T., Leying,H., Hinzpeter,M. and Karl,G.
```

```
  TITLE     Method of detecting telomerase activity
  JOURNAL   Patent: US 6221584-A 12 24-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT       13 a     16 c      9 g      6 t
ORIGIN

  Query Match             67.8%;  Score 15.6;  DB 6;  Length 44;
  Best Local Similarity   63.6%;  Pred. No. 4e+03;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 TACTGCTCAGAGUUAGGGUUAG 23
              | ||| | || |::||||::||
Db         28 TTCTGGTTAGGGTTAGGGTTAG 7


RESULT 12
AF427996/c
LOCUS       AF427996                196 bp    DNA     linear   VRT 01-NOV-2001
DEFINITION  Fringilla coelebs clone pGS-1 satellite sequence.
ACCESSION   AF427996
VERSION     AF427996.1  GI:16566655
KEYWORDS    .
SOURCE      Fringilla coelebs (chaffinch)
  ORGANISM  Fringilla coelebs
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
            Fringillidae; Fringillinae; Fringilla.
REFERENCE   1  (bases 1 to 196)
  AUTHORS   Liangouzov,I.A., Derjusheva,S.E., Saifitdinova,A.F., Malykh,A.G.
            and Gaginskaya,E.R.
  TITLE     Tandem repeat from chaffinch genome associated with interstitial
            heterochromatin
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196)
  AUTHORS   Liangouzov,I.A., Derjusheva,S.E., Saifitdinova,A.F., Malykh,A.G.
            and Gaginskaya,E.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2001) Laboratory of Chromosome Structure and
            Function, Biological Institute of Saint-Petersburg State
            University, Stary Petergof, Oranienbaumskoye sh., 2,
            Saint-Petersburg 198904, Russia
FEATURES             Location/Qualifiers
     source          1. .196
                     /organism="Fringilla coelebs"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:37598"
                     /chromosome="1; 2; 3; Ma; Mi"
                     /clone="pGS-1"
     satellite       1. .196
                     /rpt_type=tandem
BASE COUNT       40 a     55 c     59 g     42 t
ORIGIN

  Query Match             67.8%;  Score 15.6;  DB 5;  Length 196;
  Best Local Similarity   63.6%;  Pred. No. 3.3e+03;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 TACTGCTCAGAGUUAGGGUUAG 23
              | ||||  || |::||||::||
Db        101 TCCTGCCTAGGGTTAGGGTTAG 80


RESULT 13
BD084917
LOCUS       BD084917                25 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Chimeric oligonucleotides and the use thereof.
ACCESSION   BD084917
VERSION     BD084917.1  GI:22630527
KEYWORDS    JP 2001524972-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Matthes,E. and Lipinski,M.V.J.
  TITLE     Chimeric oligonucleotides and the use thereof
  JOURNAL   Patent: JP 2001524972-A 2 04-DEC-2001;
            MAX DELBRUCK CENTRUM FUR MOLEKULARE MEDIZIN
COMMENT     OS   Artificial Sequence
            PN   JP 2001524972-A/2
            PD   04-DEC-2001
            PF   04-MAY-1998 JP 1998547618
            PR   02-MAY-1997 DE 197 20 151.2
            PI   ECKART MATTHES,MARTIN VON JANTA LIPINSKI
            PC   C07H21/00
            CC   Description of Artificial Sequence:chimera oligonucleotide FH
             Key             Location/Qualifiers
            FT   source          1. .25
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        7 a      4 c      7 g      7 t
ORIGIN

  Query Match             67.0%;  Score 15.4;  DB 6;  Length 25;
  Best Local Similarity   82.4%;  Pred. No. 5.4e+03;
  Matches   14;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAG 17
              | ||||||||||::||
Db          9 GGACTGCTCAGAGTTAG 25


RESULT 14
BD084920
LOCUS       BD084920                25 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Chimeric oligonucleotides and the use thereof.
ACCESSION   BD084920
VERSION     BD084920.1  GI:22630530
KEYWORDS    JP 2001524972-A/5.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Matthes,E. and Lipinski,M.V.J.
  TITLE     Chimeric oligonucleotides and the use thereof
  JOURNAL   Patent: JP 2001524972-A 5 04-DEC-2001;
            MAX DELBRUCK CENTRUM FUR MOLEKULARE MEDIZIN
COMMENT     OS   Artificial Sequence
            PN   JP 2001524972-A/5
            PD   04-DEC-2001
            PF   04-MAY-1998 JP 1998547618
            PR   02-MAY-1997 DE 197 20 151.2
            PI   ECKART MATTHES,MARTIN VON JANTA LIPINSKI
            PC   C07H21/00
            CC   Description of Artificial Sequence:chimera oligonucleotide FH
             Key             Location/Qualifiers
            FT   source          1. .25
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        7 a      4 c      7 g      7 t
ORIGIN

  Query Match             67.0%;  Score 15.4;  DB 6;  Length 25;
  Best Local Similarity   82.4%;  Pred. No. 5.4e+03;
  Matches   14;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GTACTGCTCAGAGUUAG 17
```

```
              | ||||||||||::||
Db          9 GGACTGCTCAGAGTTAG 25


RESULT 15
BD049748/c
LOCUS       BD049748                145 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD049748
VERSION     BD049748.1  GI:22591490
KEYWORDS    JP 2001269182-A/25994.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 145)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 25994 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/25994
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .145
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT       39 a     33 c     19 g     54 t
ORIGIN

  Query Match            66.1%;  Score 15.2;  DB 6;  Length 145;
  Best Local Similarity  75.0%;  Pred. No. 5.7e+03;
  Matches   15;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 ACTGCTCAGAGUUAGGGUUA 22
              |||| ||||||: ||||: |
Db        124 ACTGGTCAGAGTCAGGGTGA 105


Search completed: November  8, 2003, 05:21:59
Job time : 1537 secs
```